Multimodal single-cell profiling reveals neuronal vulnerability and pathological cell states in focal cortical dysplasia

Isabella C. Galvão[1,2], Manuela Lemoine[1,2], Lauana A. Messias[1,2], Patrícia A.O.R.A. Araújo[1,2], Jaqueline C. Geraldis[1,2], Clarissa L. Yasuda[2,3], Marina K. M. Alvim[2,3], Enrico Ghizoni[2,3], Helder Tedeschi[2,3], Fernando Cendes[2,3], Fabio Rogerio[2,4], Iscia Lopes-Cendes[1,2], and Diogo F. T. Veiga[1,2,*]

[1]Department of Translational Medicine, School of Medical Sciences, University of Campinas, Campinas (UNICAMP), Brazil

[2]The Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, Brazil

[3]Department of Neurology, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, SP, Brazil

[4]Department of Pathology, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, Brazil

[*]Corresponding author: dfveiga@unicamp.br

ABSTRACT

Focal Cortical Dysplasia (FCD) is a neurodevelopmental condition characterized by malformations of the cerebral cortex that often cause drug-resistant epilepsy. In this study, we performed multi-omics single-cell profiling to map the chromatin accessibility and transcriptome landscapes of FCD type II, generating a comprehensive multimodal single-cell dataset comprising 61,525 cells from 11 clinical samples of lesions and controls. Our findings revealed profound chromatin, transcriptomic, and cellular alterations affecting neuronal and glial cells in FCD lesions, including the selective loss of upper-layer excitatory neurons, significant expansion of oligodendrocytes and immature astrocytic populations, and a unique neuronal subpopulation harboring dysmorphic neurons. Furthermore, we uncovered activated microglia subsets, particularly in FCD IIb cases. This comprehensive study unveils neuronal and glial cell states driving FCD development and epileptogenicity, enhancing our understanding of FCD and offering new directions for targeted therapy development.

Keywords: focal cortical dysplasia; cortical malformations; single-cell sequencing; multi-omics; gene expression; chromatin accessibility; microglia; reactive astrocytes

43 **INTRODUCTION**

44 Focal Cortical Dysplasia (FCD) is a malformation of the cerebral cortex and a major cause of
45 drug-resistant epilepsy in children. Treatment of this neurological condition with current anti-
46 seizure medications is often ineffective, frequently requiring surgical removal of the affected
47 brain tissue. FCD is stratified based on neuroimaging, histopathology, and genetic
48 characteristics[1,2]. FCD type II, the most common presentation, is characterized by
49 cytoarchitectural cortical abnormalities including loss of cortical lamination, blurring of the
50 gray-white matter junction, and abnormal cellular development[2]. FCD II cortical lesions can
51 be further divided based on cytological features: FCD IIa is characterized by the presence of
52 dysmorphic neurons, while FCD IIb presents with both dysmorphic neurons and balloon cells.
53 Dysmorphic neurons are abnormal cells featuring enlarged soma and nuclei[1,2], along with
54 expression of neurofilament markers typical of immature neurons in the developing cortex. On
55 the other hand, balloon cells are characterized by large cell bodies with vitreous, opalescent,
56 and eosinophilic cytoplasm[1,2]. These cells express markers commonly associated with
57 neuroglial progenitors, including nestin and vimentin, indicating their immature nature[3]. High
58 expression of other progenitor cell markers such as c-Myc, SOX2, and Oct-4 were also
59 observed throughout FCD II lesions[4,5].

60 At the molecular level, FCD II cases display activation of the mammalian target of
61 rapamycin complex 1 (mTORC1) signaling pathway, which is implicated in cell growth,
62 increased synaptic transmission and plasticity, as well as regulation of immune system
63 responses[6–8]. Hyperactivation of mTOR signaling, which operates through the mTORC1 and
64 mTORC2 complexes, is particularly observed in abnormal cells such as dysmorphic neurons
65 and balloon cells[9]. Increased phosphorylation of the mTORC1 substrates S6K1 and S6 proteins
66 can be used as a biomarker to distinguish FCD type II lesions[4]. Genetic alterations such as

67 somatic gain-of-function mutations in mTOR activating genes and germline loss-of-function
68 mutations of mTOR inhibitors are also found in FCD II[9–11].

69 The cellular and epileptogenic mechanisms of FCD are not yet fully understood.
70 Dysmorphic neurons appear to play a pivotal role in generating epileptic discharges[11], while
71 balloon cells are considered electrically quiescent and may even have an anti-seizure effect[3].
72 A recent study showed that dysmorphic neurons derive from excitatory neurons while balloon
73 cells have astrocytic lineage[12]. We and other groups have shown that FCD IIb lesions are also
74 characterized by increased microglia and astrocyte reactivity when compared to FCD IIa or
75 healthy controls[13–15]. However, little is known regarding the cellular landscape of FCD lesions
76 and the role of the diverse cortical cell types in the disease.

77 Single-cell RNA and transposase-accessible chromatin (ATAC) sequencing have been
78 applied to resolve the cellular diversity of the human brain in healthy and disease contexts[16–18]
79 and to discover cell states associated with neurological conditions[19–21]. Recent multimodal
80 single-cell approaches can obtain multiple measurements from the same cells simultaneously,
81 enabling the integration of epigenetic and transcriptomic modalities to identify cell types
82 involved in diseases[22,23].

83 In this study, we applied multi-omics single-cell profiling to characterize the chromatin
84 accessibility and gene expression landscapes in FCD type II, generating a comprehensive
85 multimodal single-cell dataset comprising 61,525 cells from 11 clinical samples of FCD lesions
86 and controls obtained from epilepsy surgeries. Integrative analyses of this dataset unraveled
87 cellular unbalances and cell states underlying FCD pathogenesis, and highlighted neuronal
88 populations susceptible in the disease.

89 **RESULTS**

90 **High-resolution characterization of the chromatin accessibility and transcriptome**
91 **landscapes of FCD type II**

92 We performed multi-omics single-cell sequencing of cortical brain tissue from FCD IIa (n=3)
93 and IIb (n=6) lesions, as well as from non-lesion tissues selected as controls (n=2; Figure 1A).
94 Specimens were collected during epilepsy surgery to treat drug-resistant FCD. Internal controls
95 correspond to histologically normal cortical tissues where neither abnormal cells (dysmorphic
96 neurons and balloon cells) nor architectural changes were observed. Dissociated nuclei were
97 profiled using the 10X Genomics Multiome ATAC + Gene Expression assay for simultaneous
98 measurement of chromatin accessibility and gene expression in individual nuclei. Quality
99 control was performed to remove low-quality nuclei and to filter out potential doublets, and
100 Harmony[24] was applied to ATAC and RNA datasets to integrate samples from different
101 sequencing batches (Figure 1B). The Weighted Nearest Neighbor (WNN) approach[25] from
102 Seurat was used to create an integrated multimodal (ATAC + RNA) visualization (Figure 1B).
103 The resulting multimodal single-cell compedium comprised 61,525 nuclei with paired
104 chromatin and gene expression measurements. Clustering analysis identified a total of 40
105 populations (Figure S1A), which were annotated using Azimuth[25] followed by manual
106 validation using canonical marker genes to define consensus cell types. Azimuth was able to
107 assign labels with high confidence to the majority of clusters (Figure S1B), and none of the
108 clusters were exclusive to a unique sample (Figure S1C). Consensus cell type annotation was
109 performed using a broad cell types (Figure 1C) as well as a higher-resolution taxonomy
110 including neuronal subtypes and cortical layers (Figure 1D).

111 Overall, our annotation identified the various cell types expected in the human cortex
112 such as excitatory and inhibitory neurons, astrocytes, oligodendrocytes, oligodendrocytes
113 progenitor cells (OPC), microglia, endothelial cells, and vascular and leptomeningeal cells

114 (VLMC, Figure 1C). We also identified the various subtypes of excitatory neurons (EN, Figure
115 1D) including upper layers EN L2/3 (termed EN Upper), and deep layer subtypes EN L5, EN
116 L5/6, EN L6 (EN Deep), while inhibitory neurons (IN) were further classified into Vip, Sst,
117 Pvalb, Lamp5 and Sncg/Lamp5 subtypes (Figure 1D). Based on the histological assessment,
118 most of the sequenced nuclei derived from IIb lesions, followed by internal controls and IIa
119 lesions (Figure 1E). Gene expression and open chromatin accessibility of marker genes were
120 inspected to confirm cell type identity (Figures 1F-G). For instance, *AQP4* (astrocytes),
121 *SLC17A7* (EN), *GAD2* (IN), *PLP1* (oligodendrocytes), *CD86* (microglia), and *GPR17* (OPC)
122 were marker genes with concordant expression and chromatin accessibility in their
123 corresponding cell types.

124 Using signature enrichment analysis, mTORC1 signaling was found to be elevated in
125 dysplastic tissue across all cell types except VLMC and endothelia (Figure S2A), and cells with
126 the highest mTORC1 activation were identified in FCD IIb (Figure S2B). Thus, multimodal
127 single-cell profiling uncovered the heterogenous composition of FCD type II lesions and
128 recapitulated key aspects of their pathological state.

129

130 **Unraveling cellular changes in FCD lesions**

131 Next, we determined cell type abundance changes between FCD and control samples with
132 Speckle[26], using an approach that accounts for variability in cellular frequencies among
133 patients. The analysis uncovered alterations affecting several broad cell types, including
134 decreased frequency of the EN upper population and increased frequency of oligodendrocytes
135 in both FCD IIa and IIb lesions (Figure 1H). In addition, we found an expansion of the
136 microglia population specifically in IIb donors.

137 We also leveraged the hierarchical nature of the Azimuth annotation, which classifies
138 excitatory and inhibitory neurons into 116 subclasses based on their expression of markers, to

139 find out how these neuronal subclasses are affected in the disease. The Exc L2
140 *LINC00507*/*GLRA3*, the most common EN subclass identified in the cortex of FCD samples,
141 was also the most affected EN displaying both the highest number of transcriptional changes
142 and frequency reduction in lesions (Figure 2A,B). Among IN, FCD IIa displayed an increased
143 frequency of Parvalbumin-expressing (Pvalb) interneurons Inh L3 Pvalb/*SAMD13* (Figure
144 S3C); yet, this was not the subclass with most transcriptional changes (Figure 2D). On the other
145 hand, somatostatin-expressing (Sst) subclass Inh L1-2 Sst/PRRT4 and PAX6-expressing
146 subclass Inh L1 PAX6/*MIR101-1* were affected only in FCD IIb when compared to controls
147 (Figure 2C). Altogether, this data shows that cellular abundance changes affect both neuronal
148 and non-neuronal populations in FCD pathogenesis, and revealed the vulnerability of
149 excitatory neurons in the disease.
150
151 **Identification of a disease-specific population containing dysmorphic neurons**
152 During annotation, we found a heterogeneous cluster containing both excitatory neurons and
153 oligodendrocytes that we initially termed Unclassified since Azimuth could not confidently
154 assign a cell type (Figure 1C). This cluster was highly disease-specific as it contained 1,308
155 cells, being that 1,285 cells (98%) originated from FCD (Figure 3A). Despite clustering next
156 to oligodendrocytes after data integration, 54% of nuclei were classified as EN and 29% as
157 oligodendrocytes. The cluster was detected in all 9 FCD samples, denoting its reproducibility
158 across donors (cluster #15, Figure S1C). To further characterize this disease-specific (DS)
159 population, we identified its markers and performed pathway enrichment analysis. Notably, all
160 top enriched pathways were related to neurodegenerative diseases including Parkinson's,
161 amyotrophic lateral sclerosis, and Alzheimer's (Figure 3B). Also, the majority of marker genes
162 were associated with neuronal processes including oxidative phosphorylation, Ca^{2+} transport,
163 neuron development, and neurofilament subunits (Figure 3C).

164 In particular, we found two neurofilament subunits *NEFL* and *NEFM* overexpressed in
165 the DS cluster (Figure 3C). *NEFM* was exclusively expressed in the DS cluster and not detected
166 in any other cell type in either lesions or non-lesion tissues (Figure 3D). The chromatin
167 accessibility in the *NEFM* and *NEFL* loci corroborates with their active gene expression in the
168 DS cluster (Figure 3E).

169 Antibodies against neurofilaments are often used to identify dysmorphic neurons in
170 dysplastic tissue during histopathological diagnosis of FCD[2,11]. Based on this observation, we
171 used a signature enrichment analysis to inspect whether the DS cluster harbor dysmorphic
172 neurons. Specifically, we analyzed a 30-gene signature expressed in a purified dysmorphic
173 neuron population selected using laser capture microdissection in FCD tissue[27]. We found that
174 nuclei with the highest expression of this signature were overlapping the DS cluster (Figure
175 1F). Finally, to find out whether the DS cluster could be detected in healthy individuals, we
176 integrated our FCD dataset with single-cell RNA-seq from cortical samples obtained from age-
177 matched autopsy controls[17]. This analysis confirmed that the DS cluster was not present in the
178 normal healthy cortex (Figure 1G). Taken together, these data demonstrate that the DS
179 population is uniquely present in FCD lesions and comprises dysmorphic neurons expressing
180 genes associated with neurodegenerative conditions.

181

182 **Pathological microglia cell states in FCD IIb**

183 To further characterize microglia states in FCD, we performed subclustering and re-annotation
184 of the broad microglial compartment. Subclustering and marker detection with Seurat
185 uncovered 7 subpopulations with clearly distinct gene signatures (Figure S3; full list of
186 subcluster markers provided in Supplemental Table 2). Diverse microglia signatures identified
187 by single-cell sequencing have been previously reported in neurological diseases[20,21,28]. We
188 used signatures of microglia subsets derived from healthy and epileptic temporal cortex[21] to

189 annotate subclusters in FCD. Using this approach, we identified three distinct microglia states
190 in FCD lesions (Figure 4A): a homeostatic subset (the largest population corresponding to
191 subcluster 0); a subset expressing high levels of *CD74* (subcluster 1, *CD74*$^{+,}$), and a subset
192 expressing high levels of *CD83* (subclusters 2 and 5, *CD83*$^{+}$). We also identified small
193 subclusters of lymphocyte/T cells, B cells, and perivascular macrophages that were annotated
194 using classical lymphocyte markers (Figure 4A). As shown in Figure 4B, *CD74*$^{+}$ microglia
195 expressed marker genes related to antigen processing and presentation/MHC class II genes
196 (*HLA-DRA*, *HLA-DPA1*, *HLA-DPB1*, *HLA-DRB1*, *HLA-DQA1*), adaptive immune system
197 response and microglial cell activation (*CD74*, *SC1N*, *C1QA*, *C1QB*, *C1QC*), and interferon
198 signaling (*IFNGR1*, *JAK2*, *B2M*). On the other hand, *CD83*$^{+}$ microglia expressed high levels
199 of pro-inflammatory cytokines including *IL1B*, *CCL2*, *CCL3*, *CCL4*, and *TNFSF18*.
200 Interestingly, *CD83* and *CD74* expression are largely non-overlapping, indicating that these
201 surface markers can be effectively used to distinguish microglia states in FCD (Figure 4C).
202 Of note, homeostatic microglia and activated states are expanded only in FCD IIb
203 tissues (Figure 4D). Compared to controls, FCD IIb had a higher frequency of *CD74*$^{+}$ and
204 homeostatic subclusters. Compared to FCD IIa, the expansion included *CD74*$^{+}$ and *CD83*$^{+}$
205 subpopulations as well as lymphoid/T cells (Figure 4D).
206 Next, we performed differential expression analysis and identified 133 upregulated
207 genes in FCD IIb microglia compared to controls (Figure 4E). GO terms associated with
208 upregulated genes included antigen processing and presentation, MHC class II protein
209 complex, and positive regulation of cell activation (Figure 4F), in line with the expansion of
210 *CD74*$^{+}$ microglia in FCD IIb. Motif activity analysis using scATAC-seq uncovered that top
211 upregulated motifs in FCD IIb microglia included IRF8 and STAT1/STAT2 (Figure 4G),
212 which are regulators related to microglia activation. Cells with the highest IRF8 and
213 STAT1/STAT2 motif activity overlapped the *CD74*$^{+}$ microglia (Figure 4H). Taken together,

214 these data highlight the emergence of pathological microglial states is specific to FCD IIb, and
215 indicated that microglia activation in FCD IIb is mediated by the *CD74*$^{+}$ subpopulation.
216
217 **Chromatin changes and cell-type specific gene regulation in FCD**
218 Next, we aimed to find open chromatin regions associated with gene regulation in FCD. First,
219 we found the set of differentially accessible chromatin regions (DACRs), which are peaks with
220 gain or loss of chromatin accessibility, by comparing chromatin profiles across cell types in
221 FCD against controls. Overall, FCD IIa had a higher number of DACRs, especially in upper
222 and deep excitatory neurons (Figure 5A). Most of the DACRs in FCD IIa displayed a gain of
223 accessibility in the lesion (opening peaks), while DACRs in FCD IIb were characterized by a
224 loss of chromatin accessibility (Figure 5A). Further analysis using Encode annotation found
225 that most DACRs were located within enhancer regions, with 34% overlapping proximal
226 enhancers near a transcriptional start site and 44% overlapping distal enhancers (Figure 5B).
227 We then sought to understand whether these DACRs were driving gene regulation. To
228 achieve this, we implemented a machine-learning approach based on random forests to learn
229 regulatory links using the paired single-cell chromatin and gene expression readouts in the
230 dataset (see Methods, Figure 5C). The method revealed open chromatin regions most likely to
231 predict gene expression in each cell type. For instance, we identified regulatory links for *FKBP*
232 prolyl isomerase 5 (*FKBP5*), a gene within the mTORC1 pathway upregulated in FCD
233 astrocytes (Figure 5D). The regulatory links involved DACRs with gain of chromatin
234 accessibility in FCD (regions highlighted in Figure 5D), suggesting that the increased *FKBP5*
235 expression is accompanied by chromatin opening. These peaks are also putative *FKBP5* cis-
236 regulatory elements since they are located in regions with enhancer activity according to
237 Encode. Similarly, we also found opening chromatin peaks located in enhancers linked to
238 *C1QA* and *C1QB*, which were genes involved in microglia cell activation increased in FCD IIb

239 (Figure S4). Taken together, these analyses demonstrated that chromatin changes affecting
240 enhancer regions are likely driving cell type-specific gene regulation in FCD.
241
242 **Immature astrocyte cell states and balloon cells in FCD**
243 We explored whether the astrocytic population is affected in FCD by employing an
244 unsupervised trajectory analysis. The trajectory organized astrocytes according to their
245 differentiation status, with immature and mature astrocytes on opposite ends of the trajectory
246 (Figure 6A). The root of the trajectory was chosen based on the expression of immature
247 astrocyte markers (Figure 6B). Remarkably, the quantification of FCD and control cells along
248 the trajectory revealed an expansion of immature astrocytes accompanied by a decrease of
249 mature astrocytes in FCD tissues (Figure 6C). On the other hand, fully mature astrocytes were
250 highly represented in controls (Figure 6C). We also analyzed the expression of gene signatures
251 corresponding to balloon cells, which are known to originate from the astrocytic lineage[27] as
252 well as reactive astrocytes often associated to neurodegenerative diseases[19]. Scores for both
253 balloon cells and reactive astrocytes signatures were more elevated in astrocytes from FCD
254 lesions and overlapped with immature states at the beginning of the trajectory (Figure 6D, top).
255 Of note, immature astrocytes identified as balloon cells were mostly exclusive to FCD IIb
256 cases, while reactive astrocytes were present in both FCD subtypes (Figure 6D, bottom). Genes
257 associated with balloon cells and reactive astrocytes including *CHI3L1*, *HSPB1*, *SERPINA3*,
258 *GFAP*, and *VIM* were up-regulated in FCD IIb astrocytes compared to FCD IIa or internal
259 controls (Figure 6E). Taken together, these analyses showed that immature astrocyte cell states,
260 including balloon cells and reactive astrocytes, are expanded in the disease, indicating that
261 normal astrocyte differentiation is impaired in FCD.
262

263 **DISCUSSION**

264 FCD is a leading cause of drug-resistant epilepsy in children and adolescents,
265 accounting for up to 50% of epilepsy surgeries in this age group[10]. This neurological condition
266 was found to be the most common diagnosis among children with drug-resistant focal epilepsy
267 requiring surgical intervention[29]. Despite advances in the genetic characterization of FCD,
268 brain surgery remains the only treatment option to reduce seizures in refractory cases. Our
269 limited understanding of the molecular mechanisms and the role of diverse brain cell types in
270 FCD epileptogenicity is a major obstacle to developing targeted treatments. To address this
271 need, we employed a cutting-edge multi-omics single-cell profiling approach to create a high-
272 resolution cellular census of the human cortex affected by FCD. We leveraged this unique
273 dataset to uncover cellular changes and coordinated chromatin and transcriptional alterations
274 in both neuronal and glial populations in response to the disease. Overall, we uncovered
275 neuronal, microglial and astrocytic pathological cell states involved the disease.

276 First, our data showed that both excitatory and inhibitory neurons undergo changes in
277 cellular abundance and transcriptional profiles during FCD development, yet neuronal loss
278 affects mostly the EN L2/3 subtype. Interestingly, EN in upper cortical layers were also found
279 to be reduced in cortical samples of patients with multiple sclerosis[30] and temporal lobe
280 epilepsy[31], highlighting their vulnerability not only in FCD but in other neurological
281 conditions.

282 In addition, we characterized a neuronal population specific to FCD lesions expressing
283 a dysmorphic neuron-signature and neurofilament genes. This neuronal population was notably
284 absent in non-lesion controls and the healthy cortex from autopsy cases. Also, the expression
285 of *NEFM*, encoding the medium polypeptide subunit of the neurofilament protein, was found
286 to be restricted to this subpopulation. Neurofilaments detected in the blood have emerged as
287 reliable diagnostic biomarkers in several neurological diseases[33]. The remarkable specificity of

288 *NEFM* expression in dysplastic cells suggests its potential as a blood biomarker for FCD,
289 prompting the need for additional research to explore its utility as a molecular diagnostic tool
290 in the disease.

291 Furthermore, we described the expansion of microglia and the emergence of
292 pathological cell states specifically in FCD IIb lesions. We identified two FCD-associated
293 microglial subpopulations with distinct signatures: a *CD74*$^{+}$ subset overexpressing antigen
294 presentation molecules (MHC class-II components), and a *CD83*$^{+}$ subset with increased
295 expression of pro-inflammatory cytokines. Previous scRNA-seq studies have identified
296 microglial populations in neurological conditions[20,21,28]. For instance, using scRNA-seq
297 followed by in situ immunohistochemistry, Olah et al[21] described the existence of nine
298 microglial subpopulations in the human cortex (prefrontal cortex of autopsy samples and
299 temporal cortex of epilepsy), including the *CD83*^{high} and *CD74*^{high} subsets. Mathys et al[34] also
300 identified a microglia subset with increased expression of *CD74* and MHCII-related genes in
301 patients with Alzheimer's disease. A *CD74* subset was also described in the pathology of
302 multiple sclerosis[35], and *CD74* has been proposed as a marker for reactive microglia due to its
303 increased expression levels in disease[36]. Although there is no consensus in the field on how to
304 annotate microglia cell states, we adopted the terminology from Olah et al since the *CD83* and
305 *CD74* molecules were identified as markers of microglia subclusters in FCD. Of note, we also
306 found higher lymphoid/T-cells infiltration is specific to FCD IIb, in agreement with previous
307 experimental findings[14].

308 There is strong emerging evidence of dysregulated neuroinflammation mediated by
309 microglia in FCD and epilepsy. Zimmer et al[14] showed that FCD IIb elicits a stronger
310 inflammatory response characterized by increased expression of the *HLA-II* and *CCL2* in
311 microglia near balloon cells. Kumar et al[37] described immune cell infiltration and increased
312 expression of pro-inflammatory genes (e.g. *IL1B*, *CCL4*, *CCL2*) in microglia from epilepsy

313 samples. Therefore, we postulate that the CD74^{+} subset (expressing antigen presentation
314 molecules) and the CD83^{+} subset (expressing pro-inflammatory cytokines) identified in this
315 study are the microglial populations mediating immune system activation and
316 neuroinflammation in FCD IIb.

317 Finally, we showed that astrocyte differentiation is impaired in FCD, and immature
318 astrocytic cell states are expanded in lesions. This shift from mature to immature astrocytes
319 can be explained in part by the appearance of reactive astrocytes and balloon cells, which are
320 abnormal cell populations typically found in the lesion microenvironment, that we
321 characterized as immature astrocytic states. Specifically, pseudotime trajectory analysis
322 revealed that reactive astrocytes resemble immature-like cells that emerge in FCD lesions,
323 supporting the view that astrogliosis induces a partial reversion of astrocyte maturation[38].
324 Similarly, we showed that balloon cells also represent immature astrocytic cell states. Balloon
325 cells were identified exclusively in FCD IIb, consistent with the histological characterization
326 of the clinical samples. Moving forward, it will be important to understand the contribution of
327 these immature astrocytes to FCD. Mature astrocytes are known to maintain neuronal
328 homeostasis and prevent excessive neuronal activity through neurotransmitter reuptake[38], and
329 their dysregulation can lead to epileptogenesis[39]. Thus, these immature astrocytic populations
330 might add to the intrinsic epileptogenicity in FCD.

331 Paired multimodal single-cell experiments enable exploration of cell-type specific gene
332 regulation by learning the relationship between chromatin accessibility and transcription in the
333 same cell[40]. We explored the integration of these paired readouts by predicting cell-type
334 specific regulatory links and identifying enhancer regions controlling key genes in the
335 mTORC1 pathway and microglia activation. Indeed, we anticipate that the high-resolution
336 multimodal dataset produced herein will be useful in future studies integrating chromatin and
337 transcriptomic data, including unraveling the impact of non-coding variants in open chromatin

338 regions in gene expression or identifying chromatin enhancers driving gene dysregulation in
339 FCD.

340 The findings presented in this study offer promising new directions for the development
341 of targeted drug treatments for FCD. By elucidating the underlying molecular and cellular
342 mechanisms common and specific to FCD IIa and IIb at single-cell resolution, this research
343 provides valuable insights that could inform the development of tailored therapeutic strategies
344 for each disease subtype. For instance, mTORC1 signaling was found to be elevated in FCD
345 in most cell types. Thus, mTOR inhibitors or antisense oligonucleotides targeting the mTOR
346 pathway hyperactivation may represent promising strategies for targeted treatment of FCD[11].
347 Gene therapies targeting the disease-specific population containing dysmorphic neurons can
348 also be explored since dysmorphic neurons are believed to play a crucial role in seizure
349 generation in FCD[3]. Further investigation into marker genes associated with this population
350 may lead to the identification of pharmacological interventions aimed at improving seizure
351 control in the disease.

352 Additionally, several rare neurological conditions lead to epilepsy[41]. These orphan
353 diseases pose significant challenges to research due to their low prevalence and unresolved
354 pathology[41]. However, the integration of single-cell datasets may provide a powerful approach
355 to uncover pathological cell states shared across various forms of rare focal epilepsies. By
356 comparing cellular profiles, it may be possible to identify common abnormal cell states,
357 molecular signatures and pathways, offering insights into disease mechanisms and potential
358 therapeutic targets applicable to multiple epilepsy subtypes[32].

359 In conclusion, our study uncovered diverse cellular abundance changes in FCD and
360 identified key neuronal, microglial and astrocytic subpopulations involved in the disease.
361 These findings improve our understanding of the condition and suggest novel avenues for
362 developing targeted treatments.

367 **STAR METHODS**

368 **Clinical samples and neuropathological diagnosis**

369 Fresh brain samples were collected from patients who underwent surgery for intractable
370 seizures after clinical diagnosis of FCD at Hospital de Clínicas, University of Campinas.
371 Surgical specimens were (i) formalin-fixed paraffin-embedded (FFPE) or (ii) snap-frozen in
372 liquid nitrogen and stored at −80°C. FFPE samples were submitted to diagnostic routine in
373 serial 4 μm-sections stained with hematoxylin and eosin (H&E) and submitted to
374 immunohistochemical reactions. For the latter protocol, the sections were exposed to
375 antibodies against NeuN (neuronal marker; 1:1,000, clone A60, Merck Millipore,
376 cat#MAB377, Temecula, CA, USA), MAP2 (neuronal marker; 1:1,000, clone M13, Thermo
377 Fisher, cat#13-1500, Waltham, MA, USA), SMI 32 (neuronal marker; 1:2,500, clone SMI 32,
378 Biolegend, cat#SMI-32R, San Diego, CA, USA), GFAP (astrocytic marker, 1:100, clone 6F2,
379 Dako/Agilent, cat#M0761, Santa Clara, CA, USA), vimentin (1:100, clone V9, e-
380 Bioscience/Thermo Fisher, cat#14-9897-82, Waltham, MA, USA), CD34 (1:50, clone QBEnd-
381 10, Dako, cat#M7165, Glostrup, Denmark) and CNPase (myeloarchitecture marker; 1:500,
382 clone 11-5 B, Millipore, cat#MAB326, Darmstadt, Germany), for 18 h at 4°C. Then, a
383 detection solution containing the secondary antibody and peroxidase (AdvanceTMHRP®,
384 Dako, cat#K4068, Glostrup, Denmark; or EnvisionTM Flex+, Dako, cat#K8002, Glostrup,
385 Denmark) was added for 30 min at 37°C. 3,3-diaminobenzidine (DAB) was used as a
386 chromogenic substrate and counterstaining was performed with hematoxylin. Negative
387 controls (without primary antibody) were run concurrently with all immunohistochemical
388 reactions.

389 Samples were classified as FCD IIa or IIb according to the guidelines of the
390 International League Against Epilepsy (ILAE)[2]. Specifically, specimens exhibiting cortical
391 dyslamination, hypertrophic and dysmorphic neurons [disoriented neurons with anomalous

392 cytoplasmic distribution of Nissl substance and accumulation of non-phosphorylated
393 neurofilament (SMI 32-positive)] without or with balloon cells (large cells with opaque
394 eosinophilic cytoplasm, vesicular nucleus and immunopositivity for vimentin) were classified
395 as FCD ILAE type IIa or IIb, respectively. Internal controls were derived from histologically
396 normal tissue where neither abnormal cells nor architectural changes were observed. Detailed
397 patient information is provided in the Supplemental Table 1. All procedures were approved by
398 the University of Campinas's Research Ethics Board, and written informed consent was
399 obtained from all patients before surgery.
400
401 **Nuclei isolation and multiome library construction**
402 The nuclei extraction protocol for human brain tissue was adapted from Jessa et al[42]. Briefly,
403 frozen tissue was dounce homogenized and incubated for 2 minutes in lysis buffer containing
404 0.1% NP-40. Then, suspension was filtered through a 30 µm cell strainer (MACS), and washed
405 in buffer containing 10% BSA to remove debris, followed by a centrifugation at 500g and 4ºC
406 for 5 minutes. Next, the suspension was carefully layered on double 29% and 50% iodixanol
407 (Sigma) cushion layers, and centrifuged at 10,000g and 4ºC for 30 minutes for additional
408 removal of debris and cell aggregates. Retrieved nuclei were resuspended and counted. Nuclei
409 were stained with Trypan blue and counted using a hemocytometer. Nuclei integrity was
410 evaluated under a microscope using 40X or 60X magnifications, and nuclei were considered
411 viable if they were roundish and with an intact membrane. Next, we proceeded to the steps of
412 the library construction, including open chromatin transposition and droplet (GEM) formation
413 using a Chromium Controller equipment (10X Genomics). Libraries were prepared using the
414 Multiome ATAC + Gene Expression kit (10X Genomics) following the manufacturer's
415 protocol CG000338 revision E. The quality of purified ATAC and RNA libraries was assessed
416 using a TapeStation equipment (Agilent Technologies).

417

418 **Sequencing and raw data processing**

419 Multiome libraries were sequenced on either Illumina HiSeqX Ten or Illumina NovaSeq 6000
420 instruments according to the sequencing depth and the read length recommended in the
421 Multiome kit protocol (CG000338). Demultiplexing, genome alignment, gene quantification,
422 and peak accessibility in single nuclei were performed with the Cell Ranger ARC pipeline[43].
423 Specifically, fastq files were generated using cellranger mkfastq, then processed using the
424 cellranger-arc count v.2.0.2 and the human genome GRCh38 to generate the count matrices for
425 the RNA and ATAC modalities. Next, we performed data processing, normalization, and batch
426 integration on both assays independently, using Seurat v.5.0.1[44] and Signac v.1.12.0[45] in the R
427 environment v.4.1.2.

428

429 **scRNA-seq analysis**

430 Samples were preprocessed in Seurat to remove low-quality nuclei using the following criteria:
431 400 < nFeatures < 7,000, 500 < UMI counts < 50,000, and percentage of mitochondrial reads
432 (percent.mt) < 15. Identification of doublets was performed per sample using scDblFinder
433 v.1.8.0[46]. Next, standard Seurat data processing and normalization steps were performed:
434 NormalizeData, FindVariableFeatures, ScaleData, RunPCA, RunUMAP, FindNeighbors, and
435 FindClusters. For each sample, the filtered counts matrix was log-normalized with regression
436 of sequencing mitochondrial gene percentages. The RunUMAP function was used with the first
437 40 PCs identified in the elbow plot to perform dimensional reduction. Clustering was
438 performed using FindNeighbors and FindClusters functions using the smart local moving
439 (SLM) algorithm, with a resolution of 1.2.

440 To correct batch effect among samples, we used Harmony v.1.2[24] to integrate sample-level
441 PCA projections using automatic hyperparameter optimization and default parameters.

442 Harmony implements an algorithm that projects cells into a common space where they cluster
443 by cell type rather than other sample attributes such as sequencing batch.
444 Finally, we used SCTransform to create a normalized expression matrix containing all samples,
445 and with regression of sequencing batch, brain region, and mitochondrial gene percentages.
446 DEGs between FCD and controls were calculated using MAST or the Wilcox test implemented
447 in Seurat as indicated. For MAST, DEGs were selected using an adjusted P value of less than
448 0.05. The Wilcox test was applied using the SCT normalized data, using a log2 fold-change
449 threshold of 0.25 and an adjusted P value of less than 0.05.
450
451 **scATAC-seq analysis**
452 Samples were preprocessed in Signac to remove low-quality nuclei using the following criteria:
453 1,000 < ATAC fragments < α(ATAC fragments), 400 < ATAC fragments in peaks < α (ATAC
454 fragments in peaks), percent.mt < 15, nucleosome signal < 2, TSS enrichment < 2. Upper limits
455 (α) of ATAC fragments per sample were computed as mean plus 2 standard deviations. Peaks
456 within each sample were identified using MACS2[47]. Standard processing and normalization
457 steps were performed as follows: RunTFIDF, FindTopFeatures, RunSVD, and FindClusters.
458 For each sample, the peak counts matrix was normalized using the RunTFIDF function, to
459 correct for differences in nuclei sequencing depth. Dimensional reduction was performed with
460 latent semantic indexing (LSI) using the RunSVD function. UMAP projection was performed
461 utilizing LSI components 2-30, as the first LSI component was confirmed to be a technical
462 variation. Before sample integration, we created a common peak set across samples using
463 MACS2 and quantified peaks in each sample using the FeatureMatrix function. Peaks in non-
464 standard chromosomes and blacklisted regions were removed from downstream analyses. We
465 then re-computed RunTFIDF, FindTopFeatures, and RunSVD on the merged object and
466 applied Harmony to integrate sample-level low-dimensional cell embeddings (LSI components

467 2-30) using the same parameters defined in the RNA modality. Differentially accessible
468 chromatin regions (DACRs) between FCD tissues and controls were calculated using
469 FindMarkers with the LR test, with a min.pct of 0.05, an adjusted P value less than 0.05, and
470 included the number of ATAC fragments as a latent variable.
471
472 **Multimodal data integration**
473 To integrate gene expression data with chromatin accessibility data, we performed Weighted
474 Nearest Neighbor (WNN) analysis[25], an unsupervised approach that learns the relative utility
475 of each modality and builds a joint graph representing a weighted combination of RNA and
476 ATAC modalities. WNN was performed using the FindMultiModalNeighbors function with
477 20 neighbors (k.nn), based on the UMAP reductions obtained after Harmony integration. The
478 WNN graph was used for creating a joint UMAP visualization and clustering based on the
479 WNN graph was performed with the FindClusters function using the smart local moving
480 (SLM) algorithm. WNN clusters with fewer than 50 nuclei, a high percentage of doublets, or
481 enriched in mitochondrial markers were removed from downstream analyses.
482
483 **Cluster annotation**
484 We used the Azimuth R package[25] to perform cell type annotation in the FCD dataset, using as
485 a reference the Allen atlas of the human cortex[48]. Azimuth performed label-transferring by
486 assigning to each cell in the FCD dataset the most similar cell type in the reference based on
487 gene expression similarity. The assigned cell type chosen for each cluster was defined as the
488 cell type with the highest mean frequency within the cluster. The Azimuth annotation was
489 manually verified using the marker genes shown in Figure 1C. Cluster and subcluster markers
490 were calculated using the function FindAllMarkers based on the SCT normalized data, with

491 the Wilcox test using a log2 fold-change threshold of 0.25 and an adjusted P value of less than
492 0.05.
493
494 **Differential cellular abundance**
495 We used the function propeller of the Speckle R package[26] to detect changes in cellular
496 composition between FCD tissues and controls. The propeller function calculated cell type
497 proportions in biological replicates, and fitted a linear model for each cell type with the tissue
498 condition and brain cortical region as covariates.
499
500 **Integration with scRNA-seq from normal cortex**
501 We obtained normal cortex scRNA-seq data from three autopsy donors with no history of
502 neurological disease from Silleti et al[17]. To reduce computational costs, we randomly sampled
503 10,000 neuronal and 10,000 non-neuronal cortical nuclei from these donors, and processed this
504 data along with our FCD dataset following the procedures described in the "scRNA-seq
505 analysis" section. We used Harmony v.1.2 to integrate all samples with automatic
506 hyperparameter optimization and default parameters. Clustering was performed with
507 FindNeighbors and FindClusters functions using the smart local moving (SLM) algorithm and
508 a resolution of 2.0.
509
510 **Gene signature analysis**
511 We used the AddModuleScore function from Seurat and the AUCell[49] R package to calculate
512 gene set enrichment at the nuclei level. The AddModuleScore function from Seurat was
513 calculated based on the SCT-normalized RNA data. AUCell computes the gene set enrichment
514 scores based on ranked gene lists (highest to lowest gene expression) in each cell. AUCell

515 scores were computed using the calcAUC function using the top 5% of genes. The obtained
516 threshold was used to classify cells with active status for the given signature.
517
518 **Motif enrichment analysis**
519 We performed motif analysis using chromVAR v. 3.3.2[50] and Signac. Motif activity at the
520 nuclei level was calculated using the RunChromVAR function in Signac after matching the set
521 of background peaks. Differential motif activity was computed using FindAllMarkers and
522 FindMarkers (pairwise comparisons) using the LR test, with a min.pct of 0.05, an adjusted P
523 value less than 0.05, and included the number of ATAC fragments as a latent variable. We also
524 filtered context-relevant motifs using the regulator expression levels. The regulator associated
525 with the motif needed to be expressed in at least 10% of the cells in the cluster.
526
527 **Trajectory Inference Analysis**
528 Pseudotime trajectory analysis was conducted using Monocle v3[51,52]. Briefly, RNA counts
529 were preprocessed using the preprocess_cds function with the PCA method. To correct batch
530 effects, we used the align_cds function[53] with the sample as an alignment group and a residual
531 formula including the sequencing batch, brain region, and mitochondrial genes percentage as
532 covariates. Nuclei were then projected onto the UMAP space[54] and divided into clusters and
533 partitions with the cluster_cells function[55]. Subsequently, an unsupervised trajectory was
534 identified using the SimplePPT (Probabilistic Pseudotime Trajectory) algorithm implemented
535 in the learn_graph function[56], with the root set according to the expression of immature cellular
536 markers. Pseudotime values were automatically assigned to individual cells based on their
537 distance along the trajectory relative to the root node. Genes with differential expression along
538 the trajectory were calculated using the graph_test function, with adjusted P-value less than
539 0.05 and arranged according to Moran's I statistic. For each of these genes, differential

540 expression between tissue conditions was assessed using the fit_models and coefficient_table
541 functions, using a regression test with an adjusted P value less than 0.05, minimum fold-change
542 of 1.5, and with the number of counts, sequencing batch, brain region and mitochondrial gene
543 percentages as covariates.
544
545 **Inference of regulatory links**
546
547 We used a regression approach based on random forests to find open chromatin regions
548 associated with gene regulation. The goal was to predict the expression levels of the target gene
549 based on the accessibility of ATAC peaks (i.e. cis-regulatory elements) around the gene
550 promoter. Considering that a target gene *g* has *n* chromatin peaks, the linear regression was
551 defined as:
552

553 $$x_{gi} = \sum_{i=1}^{n} w_j a_{ji} + \epsilon_i$$

554
555 where x_{gi} is the expression of *g* in cell *i*, a_{ji} is the accessibility of a given peak *j* in cell *i*, w_j is
556 the weight and ϵ_i is the noise term. For the regression, we tested all peaks located in a 500kb
557 window around the gene's transcriptional start site (TSS), to include both proximal and distal
558 regulatory elements. The optimal weights, which were associated with the importance of the
559 cis-regulatory regions as predictors of gene expression, were determined using random forests.
560 Specifically, the regression was implemented using the GENIE3 R package[49], setting the
561 number of trees to 200, and peaks with weights in the top 5% (95% percentile) for each gene
562 were selected. We used the Encode annotation[57] downloaded from https:
563 //screen.encodeproject.org to identify open chromatin regions with potential enhancer activity.
564 Encode enhancers are classified based on their relative location to the TSS as proximal

565 enhancer-like sequences (pELS) if located within 200bp < TSS ≤ 2kb, or distal enhancer-like
566 sequence (dELS) when located more than 2kb from a TSS. We used both pELS and dELS
567 datasets for annotation.
568
569

570 **REFERENCES**

571 1. Blümcke, I. *et al.* The clinicopathologic spectrum of focal cortical dysplasias: A consensus
572 classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods
573 Commission. *Epilepsia* **52**, 158–174 (2011).
574 2. Najm, I. *et al.* The ILAE consensus classification of focal cortical dysplasia: An update
575 proposed by an ad hoc task force of the ILAE diagnostic methods commission. *Epilepsia*
576 **63**, 1899–1919 (2022).
577 3. Ii, P. H. I. & Crino, P. B. Focal Cortical Dysplasia: Gene Mutations, Cell Signaling, and
578 Therapeutic Implications. *Annual Review of Pathology: Mechanisms of Disease* **12**, 547–
579 571 (2017).
580 4. Orlova, K. A. *et al.* Early Progenitor Cell Marker Expression Distinguishes Type II From
581 Type I Focal Cortical Dysplasias. *J Neuropathol Exp Neurol* **69**, 850–863 (2010).
582 5. Hanai, S. *et al.* Abnormal maturation of non-dysmorphic neurons in focal cortical
583 dysplasia: Immunohistochemical considerations. *Seizure* **19**, 274–279 (2010).
584 6. Lasarge, C. L. & Danzer, S. C. Mechanisms regulating neuronal excitability and seizure
585 development following mTOR pathway hyperactivation. *Frontiers in Molecular*
586 *Neuroscience* **7**, (2014).
587 7. Powell, J. D., Pollizzi, K. N., Heikamp, E. B. & Horton, M. R. Regulation of Immune
588 Responses by mTOR. *Annual Review of Immunology* **30**, 39–68 (2012).
589 8. Weichhart, T., Hengstschläger, M. & Linke, M. Regulation of innate immune cell function
590 by mTOR. *Nat Rev Immunol* **15**, 599–614 (2015).
591 9. Lin, Y. *et al.* Similar PDK1–AKT–mTOR pathway activation in balloon cells and
592 dysmorphic neurons of type II focal cortical dysplasia with refractory epilepsy. *Epilepsy*
593 *Research* **112**, 137–149 (2015).

594 10. Lim, J. S. *et al.* Brain somatic mutations in MTOR cause focal cortical dysplasia type II
595 leading to intractable epilepsy. *Nat Med* **21**, 395–400 (2015).
596 11. Gerasimenko, A., Baldassari, S. & Baulac, S. mTOR pathway: Insights into an established
597 pathway for brain mosaicism in epilepsy. *Neurobiology of Disease* **182**, 106144 (2023).
598 12. Baulac, S. *et al. Single-Cell Genotyping and Transcriptomic Profiling in Focal Cortical*
599 *Dysplasia*. https://www.researchsquare.com/article/rs-4014535/v1 (2024)
600 doi:10.21203/rs.3.rs-4014535/v1.
601 13. Boer, K. *et al.* Evidence of activated microglia in focal cortical dysplasia. *Journal of*
602 *Neuroimmunology* **173**, 188–195 (2006).
603 14. Zimmer, T. S. *et al.* Balloon cells promote immune system activation in focal cortical
604 dysplasia type 2b. *Neuropathol Appl Neurobiol* **47**, 826–839 (2021).
605 15. Galvão, I. C. *et al.* Identifying cellular markers of focal cortical dysplasia type II with cell-
606 type deconvolution and single-cell signatures. *Sci Rep* **13**, 13321 (2023).
607 16. Bakken, T. E. *et al.* Comparative cellular analysis of motor cortex in human, marmoset
608 and mouse. *Nature* **598**, 111–119 (2021).
609 17. Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain.
610 *Science* **382**, eadd7046 (2023).
611 18. Walter, T. J., Suter, R. K. & Ayad, N. G. An overview of human single-cell RNA sequencing
612 studies in neurobiological disease. *Neurobiology of Disease* **184**, 106201 (2023).
613 19. Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia.
614 *Nature* **541**, 481–487 (2017).
615 20. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development
616 of Alzheimer's Disease. *Cell* **169**, 1276-1290.e17 (2017).

617 21. Olah, M. *et al.* Single cell RNA sequencing of human microglia uncovers a subset
618 associated with Alzheimer's disease. *Nat Commun* **11**, 6129 (2020).
619 22. Flynn, E., Almonte-Loya, A. & Fragiadakis, G. K. Single-Cell Multiomics. *Annual Review of*
620 *Biomedical Data Science* **6**, 313–337 (2023).
621 23. Armand, E. J., Li, J., Xie, F., Luo, C. & Mukamel, E. A. Single-Cell Sequencing of Brain Cell
622 Transcriptomes and Epigenomes. *Neuron* vol. 109 11–26 (2021).
623 24. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with
624 Harmony. *Nat Methods* **16**, 1289–1296 (2019).
625 25. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29
626 (2021).
627 26. Phipson, B. *et al.* propeller: testing for differences in cell type proportions in single cell
628 data. *Bioinformatics* **38**, 4720–4726 (2022).
629 27. Baulac, S. *et al.* Single-cell genotyping and transcriptomic profiling in focal cortical
630 dysplasia. Preprint at https://doi.org/10.21203/rs.3.rs-4014535/v1 (2024).
631 28. Paolicelli, R. C. *et al.* Microglia states and nomenclature: A field at its crossroads. *Neuron*
632 **110**, 3458–3483 (2022).
633 29. Blumcke, I. *et al.* Histopathological Findings in Brain Tissue Obtained during Epilepsy
634 Surgery. *N Engl J Med* **377**, 1648–1656 (2017).
635 30. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis.
636 *Nature* **573**, 75–82 (2019).
637 31. Pfisterer, U. *et al.* Identification of epilepsy-associated neuronal subtypes and gene
638 expression underlying epileptogenesis. *Nat Commun* **11**, 5038 (2020).
639 32. Sizemore, G. *et al.* Temporal Lobe Epilepsy, Stroke, and Traumatic Brain Injury:
640 Mechanisms of Hyperpolarized, Depolarized, and Flow-Through Ion Channels Utilized as

641 Tri-Coordinate Biomarkers of Electrophysiologic Dysfunction. *OBM Neurobiol* **2**, 009
642 (2018).
643 33. Khalil, M. *et al.* Neurofilaments as biomarkers in neurological disorders — towards
644 clinical application. *Nat Rev Neurol* **20**, 269–287 (2024).
645 34. Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**,
646 332–337 (2019).
647 35. Masuda, T. *et al.* Spatial and temporal heterogeneity of mouse and human microglia at
648 single-cell resolution. *Nature* **566**, 388–392 (2019).
649 36. Potru, P. S. & Spittau, B. CD74: a prospective marker for reactive microglia? *Neural*
650 *Regen Res* **18**, 2673–2674 (2023).
651 37. Kumar, P. *et al.* Single-cell transcriptomics and surface epitope detection in human brain
652 epileptic lesions identifies pro-inflammatory signaling. *Nat Neurosci* **25**, 956–966 (2022).
653 38. Lattke, M. & Guillemot, F. Understanding astrocyte differentiation: Clinical relevance,
654 technical challenges, and new opportunities in the omics era. *WIREs Mech Dis* **14**, e1557
655 (2022).
656 39. Verhoog, Q. P., Holtman, L., Aronica, E. & van Vliet, E. A. Astrocytes as Guardians of
657 Neuronal Excitability: Mechanisms Underlying Epileptogenesis. *Front. Neurol.* **11**, (2020).
658 40. Baysoy, A., Bai, Z., Satija, R. & Fan, R. The technological landscape and applications of
659 single-cell multi-omics. *Nat Rev Mol Cell Biol* **24**, 695–713 (2023).
660 41. Kioutchoukova, I. P. *et al.* Neurologic orphan diseases: Emerging innovations and role
661 for genetic treatments. *World J Exp Med* **13**, 59–74 (2023).
662 42. Jessa, S. *et al.* Stalled developmental programs at the root of pediatric brain tumors. *Nat*
663 *Genet* **51**, 1702–1713 (2019).

664 43. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat*
665 *Commun* **8**, 14049 (2017).
666 44. Dictionary learning for integrative, multimodal and scalable single-cell analysis | Nature
667 Biotechnology. https://www.nature.com/articles/s41587-023-01767-y.
668 45. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state
669 analysis with Signac. *Nat Methods* **18**, 1333–1341 (2021).
670 46. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet
671 identification in single-cell sequencing data using scDblFinder. *F1000Res* **10**, 979 (2021).
672 47. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137
673 (2008).
674 48. Comparative cellular analysis of motor cortex in human, marmoset and mouse | Nature.
675 https://www.nature.com/articles/s41586-021-03465-8.
676 49. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat*
677 *Methods* **14**, 1083–1086 (2017).
678 50. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring
679 transcription-factor-associated accessibility from single-cell epigenomic data. *Nature*
680 *Methods* **14**, 975–978 (2017).
681 51. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by
682 pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–386 (2014).
683 52. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis.
684 *Nature* **566**, 496–502 (2019).
685 53. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell
686 RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat*
687 *Biotechnol* **36**, 421–427 (2018).

688 54. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and
689 Projection for Dimension Reduction. Preprint at http://arxiv.org/abs/1802.03426 (2020).
690 55. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-
691 connected communities. *Sci Rep* **9**, 5233 (2019).
692 56. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat*
693 *Methods* **14**, 979–982 (2017).
694 57. Moore, J. E. *et al.* Expanded encyclopaedias of DNA elements in the human and mouse
695 genomes. *Nature* **583**, 699–710 (2020).
696
697

Acknowledgments

We thank the patients and their families who participated in this study. We thank Claudia Kleimann, Steven Hebert, Bhavyaa Chandarana and Tomas V. Waichman for advice on single-cell analysis; Damien Faury for sharing the nuclei extraction protocol; Danielle C. Bruno for discussions regarding the tissue dissociation. This work was supported by the following grants: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Young Investigator grant 2019/07382-2 (D.F.T.V); the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation grant DAF2021-237598 (D.F.T.V.); FAPESP doctoral fellowship 2022/01530-2 (I.C.G); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) master fellowship (L.A.M.); Conselho Nacional de Pesquisa (CNPq) grant 311923/2019-4 (I.L.-C.); FAPESP doctoral fellowship 2020/06168-4 (J.C.G.), FAPESP regular grant 2019/08259-0 (F.R.), and FAPESP Research Innovation and Dissemination Center grant 2013/07559-3 (F.C.).

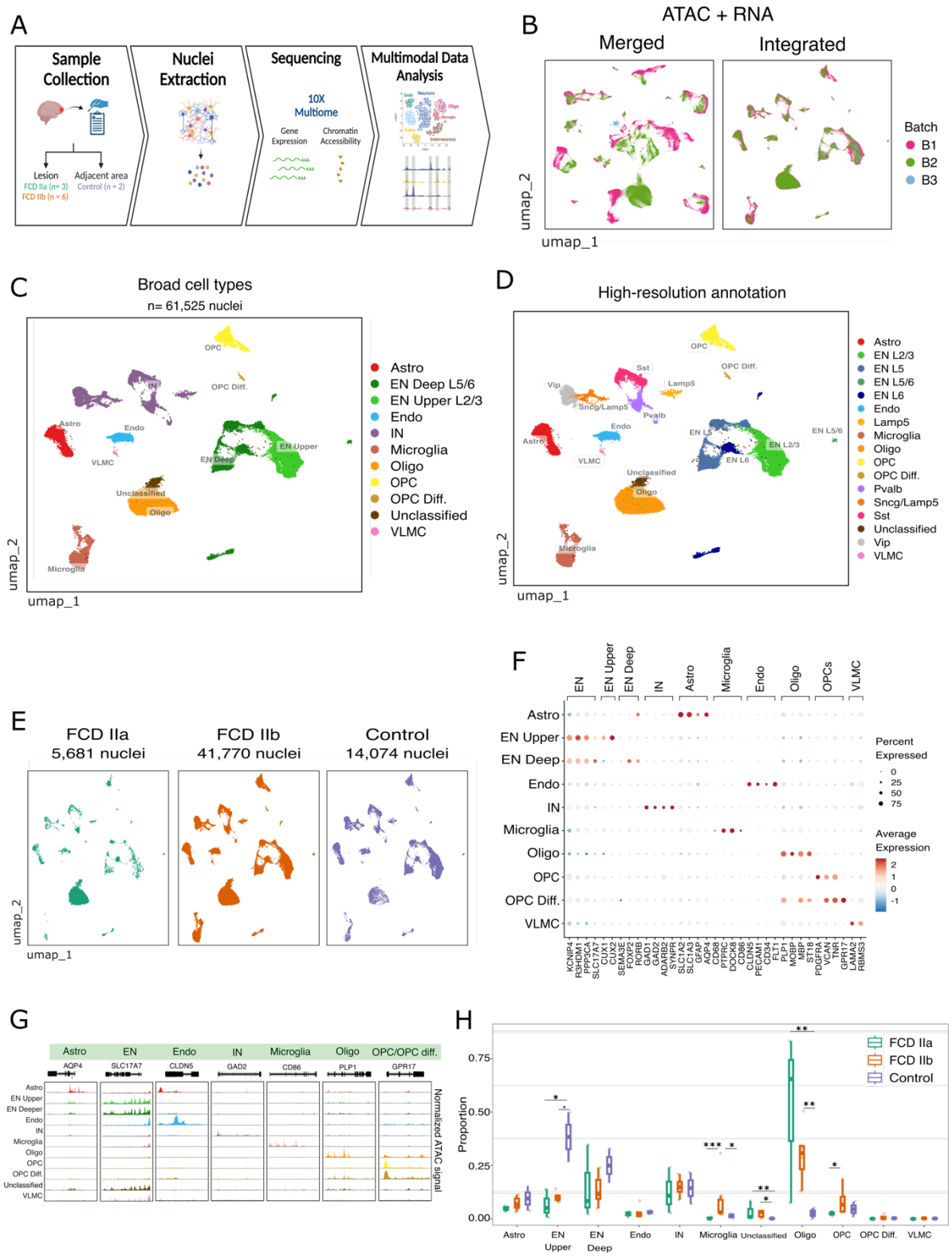
Author contributions

I.C.G. and D.F.T.V. designed the study. L.A.M. and M.L. performed single-cell assays. I.C.G. and D.F.T.V. performed data analyses. P. A., J.C.G., C.L.Y., E.G., H.T., M.K.M.A., J.C.G, F.C., and I.L-C contributed to clinical sample collection. F.R. performed histopathological analyses. I.C.G. and D.F.T.V. wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

723 **FIGURES**



724

725

726 **Figure 1. Overview of the multimodal single-cell sequencing in FCD type II.**

727

(A) Schematic of FCD samples and study design. Cortical tissues from lesions and adjacent non-lesion areas in individuals with FCD IIa and IIb were profiled by multiome single-cell ATAC and RNA sequencing. Created with Biorender.com.

(B) Uniform manifold approximation and projection (UMAP) visualization of joint modalities (ATAC + RNA) after sample integration. Nuclei are colored by sequencing batch.

(C) UMAP joint visualization of nuclei colored by broad cell type annotation.

(D) UMAP joint visualization colored using a high-resolution annotation with classification of neuronal subtypes.

(E) UMAP plots depicting nuclei by tissue condition.

(F) Dot plot showing gene expression of known marker genes for major cortical cell types. Dot size corresponds to the fraction of cells expressing the marker, and color corresponds to the expression levels.

(G) Coverage plots of marker genes depicting chromatin accessibility across cell types. The normalized ATAC signal is depicted in a region of +/- 1kb around gene coordinates.

(H) Box plots showing cell type proportion by tissue condition. In box plots, the median is indicated by the center line; box limits represent upper and lower quartiles; and whiskers extend to 1.5 times the interquartile range. Overlay dots represent cell type proportions in individual samples. Significant changes in cell types were detected using Speckle. ˙P < 0.1, *P < 0.5, **P < 0.01, ***P < 0.001.

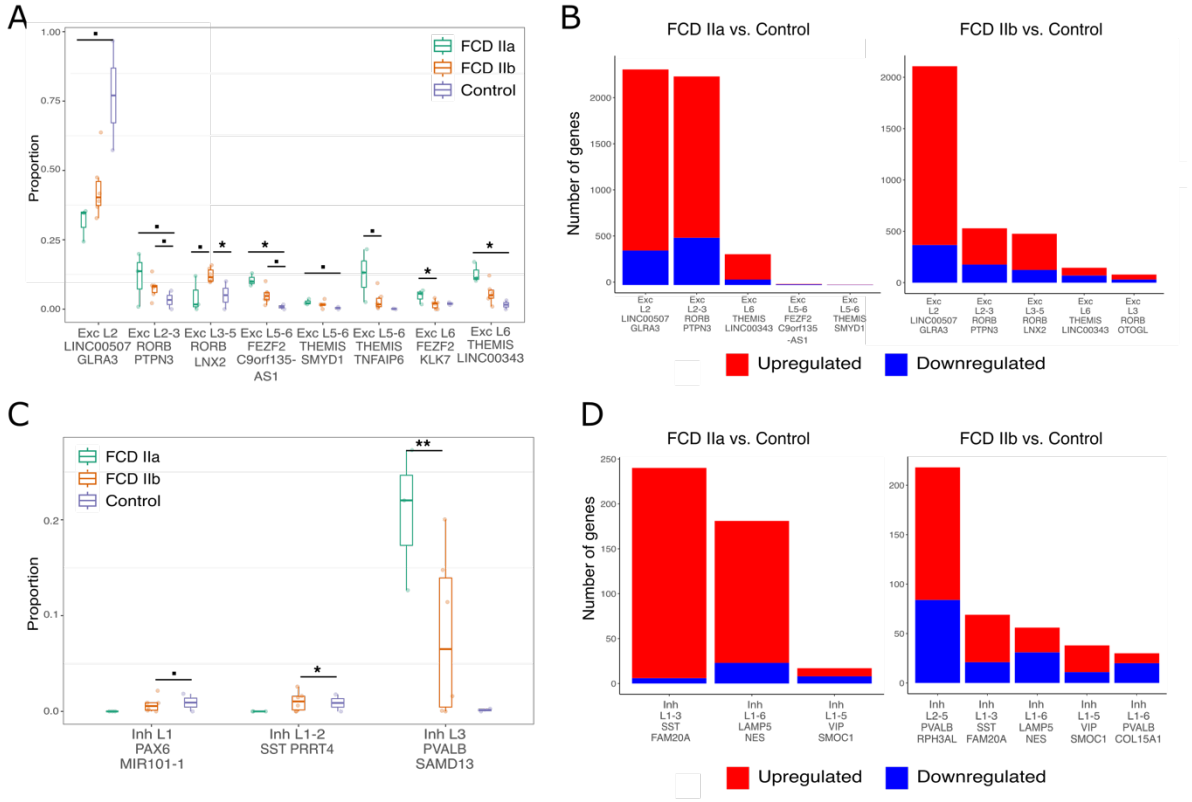


Figure 2. Excitatory and inhibitory neuron subtypes affected in FCD.

(A) Box plots showing excitatory neurons (EN) subtype frequency in FCD and controls. In box plots, the median is indicated by the center line; box limits represent upper and lower quartiles; and whiskers extend to 1.5 times the interquartile range. Overlay dots represent proportions in individual samples. Significant changes were detected using Speckle. ·P < 0.1, *P < 0.5, **P < 0.01, ***P < 0.001.

(B) Bar graphs depicting the number of differentially expressed genes (DEGs) in EN subtypes. DEGs between FCD and controls were computed using MAST with adj. P < 0.05. Only the five subtypes with the most transcriptional changes are shown.

(C) Box plots showing inhibitory neurons (IN) subtype frequency in FCD and controls. Box plots legends and significance levels are defined as in (A).

(D) Bar graphs depicting DEGs in IN subtypes. DEGs between FCD and controls were computed using MAST with adj. P < 0.05. Only the five subtypes with the most transcriptional changes are shown.

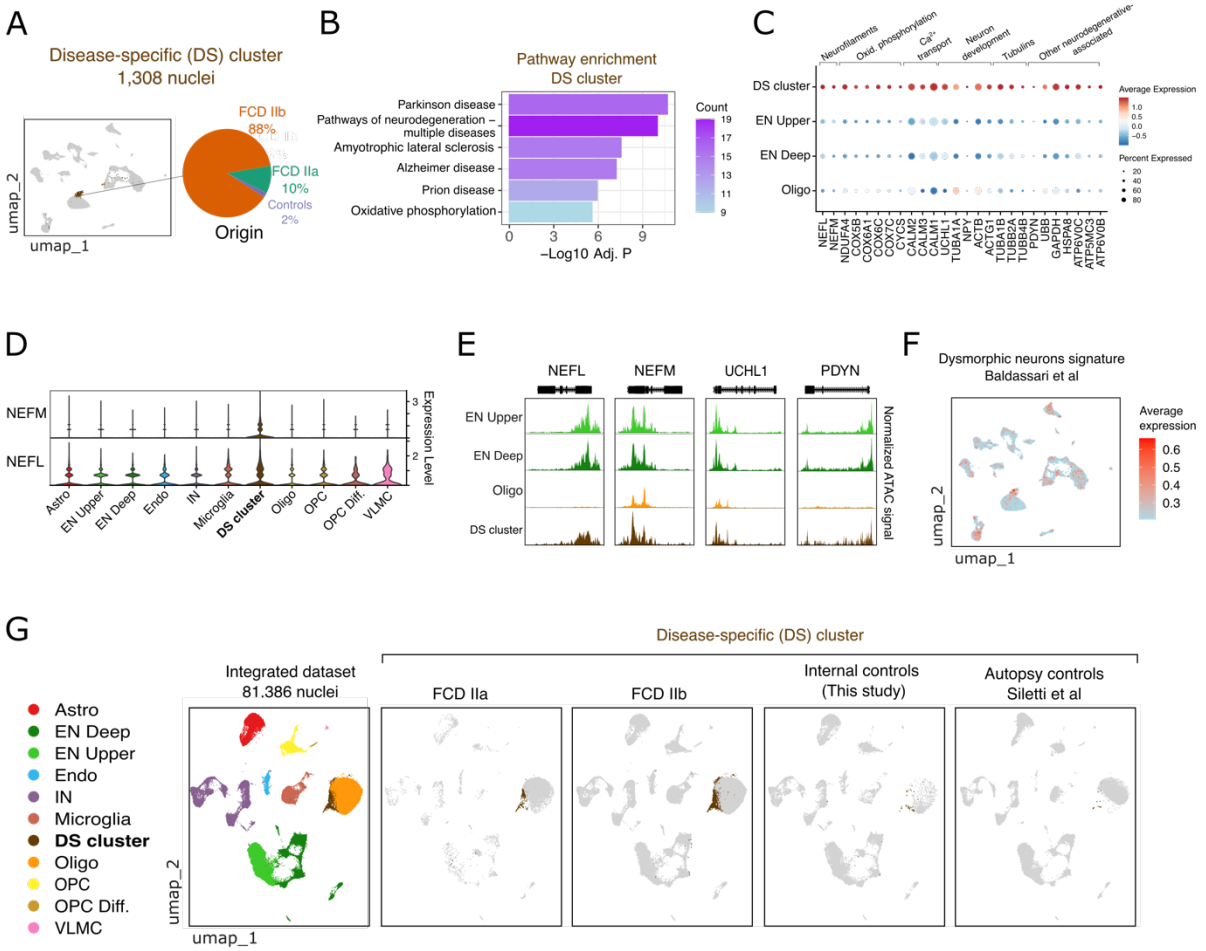


Figure 3. Characterization of a disease-specific cluster containing dysmorphic neurons.

(A) UMAP highlighting the disease-specific (DS) cluster. The pie chart indicates the percentage of nuclei from FCD IIa, IIb, and controls within the cluster.

(B) Pathway enrichment of DS cluster markers. Marker genes were identified using Seurat´s FindMarkers and KEGG pathway enrichment was performed using g:Profiler. The bars represent the significance score of pathway enrichment, and the color indicates the number of genes in each pathway.

(C) Dot plot denoting expression of marker genes associated with neurons in the DS cluster. Neuronal genes are grouped by biological function. Dot size and color correspond to expression frequency and level in the cluster, respectively.

(D) Violin plots denoting gene expression of *NEFM* and *NEFL* neurofilaments in the various cell types.

(E) Coverage plots of neuronal genes expressed in the DS cluster. The normalized ATAC signal is depicted in a region of +/- 1kb around gene coordinates.

798 **(F)** UMAP showing the average expression of a dysmorphic neuron (DN) signature obtained
799 from Baldassari et al. Signature expression was calculated using the AddModuleScore function
800 from Seurat.
801
802 **(G)** UMAP RNA plots showing the DS cluster population in FCD IIa and IIb lesions,
803 histologically normal tissue from FCD individuals (internal controls), and healthy cortex
804 (autopsy controls). The leftmost UMAP depicts annotated cell types in the integrated dataset,
805 which was created by integrating scRNA-seq data from this study with Siletti et al using
806 Harmony.
807
808
809

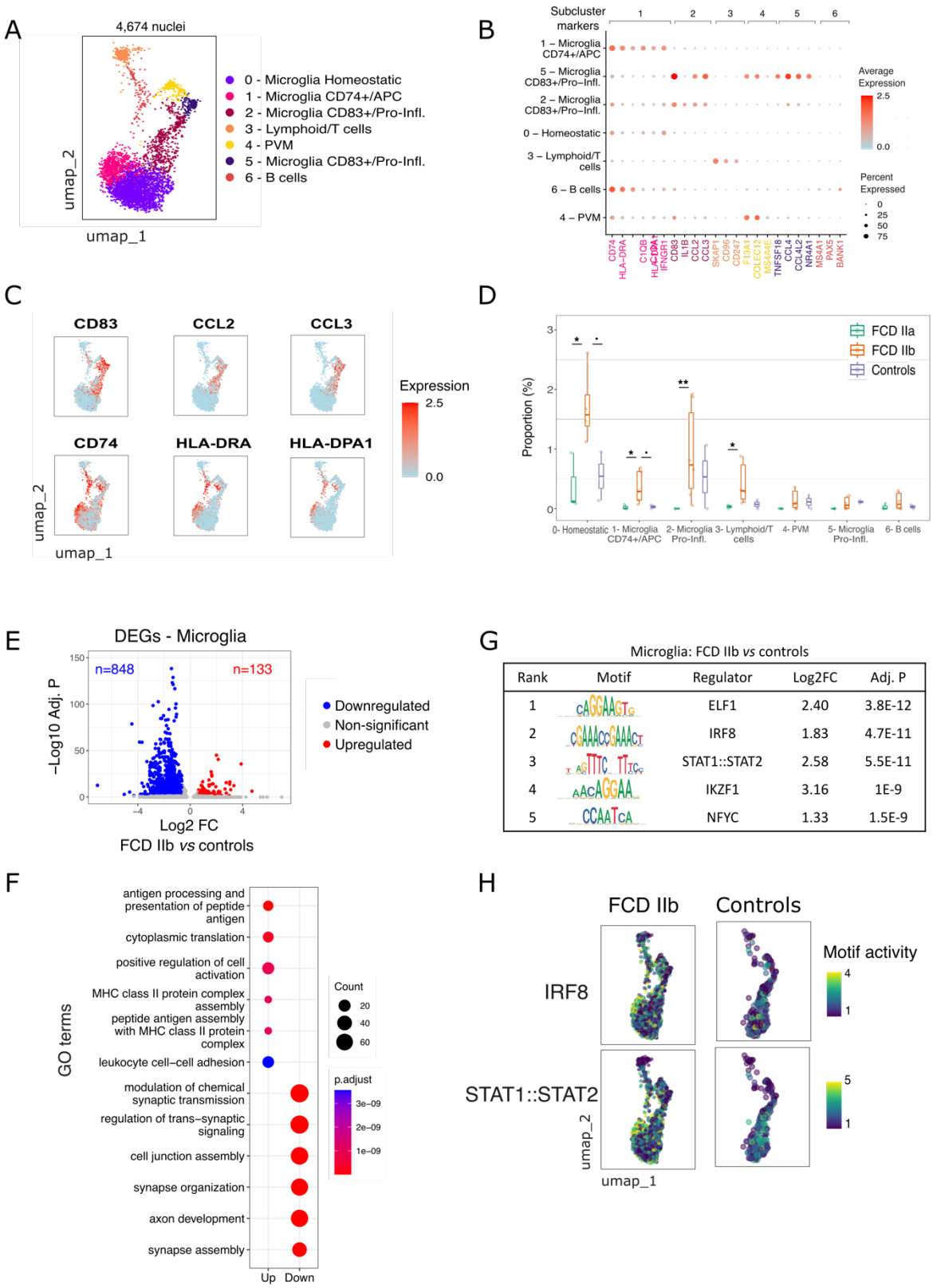


Figure 4 – Microglia pathological cell states and activation in FCD type IIb.

(A) UMAP visualization of microglia and lymphoid subclusters sequenced in FCD lesions.

(B) Feature plots depicting expression of key genes distinguishing microglia subsets *CD83*$^{+}$ (*CD83*, *CCL2*, *CCL3*) and *CD74*$^{+}$ (*CD74*, *HLA-DRA*, *HLA-DPA1*).

817
818 **(C)** Dot plot denoting the expression of key marker genes in microglia and lymphoid
819 subclusters.
820
821 **(D)** Box plots denoting subcluster frequency by tissue condition. Significant changes were
822 detected using Speckle. Box plots and significance legends are defined in Fig. 1H.
823
824 **(E)** Volcano plot depicting differentially expressed genes (DEGs) in FCD IIb microglia. DEGs
825 were obtained using a Wilcox test after SCT transformation with correction for sequencing
826 batch, mitochondrial gene percentage, and brain region. The *x*-axis is the fold-change (log2)
827 expression in FCD IIb *vs* controls, and the *y*-axis depicts the significance of the change (-log10
828 of the adjusted P-value). The color indicates the gene status.
829
830 **(F)** Dot plots showing top enriched GO terms in up- and down-regulated genes from (E). Dot
831 size and color correspond to the number of genes associated with the GO term and enrichment
832 P-value, respectively.
833
834 **(G)** Top five differential motifs in microglia from FCD IIb versus controls, as measured by
835 chromVAR (see Methods).
836
837 **(H)** Feature plots depicting IRF8 and STAT1/STAT2 motif activity in microglia open
838 chromatin regions from FCD IIb and controls. Cells are colored by the chromVAR score.
839

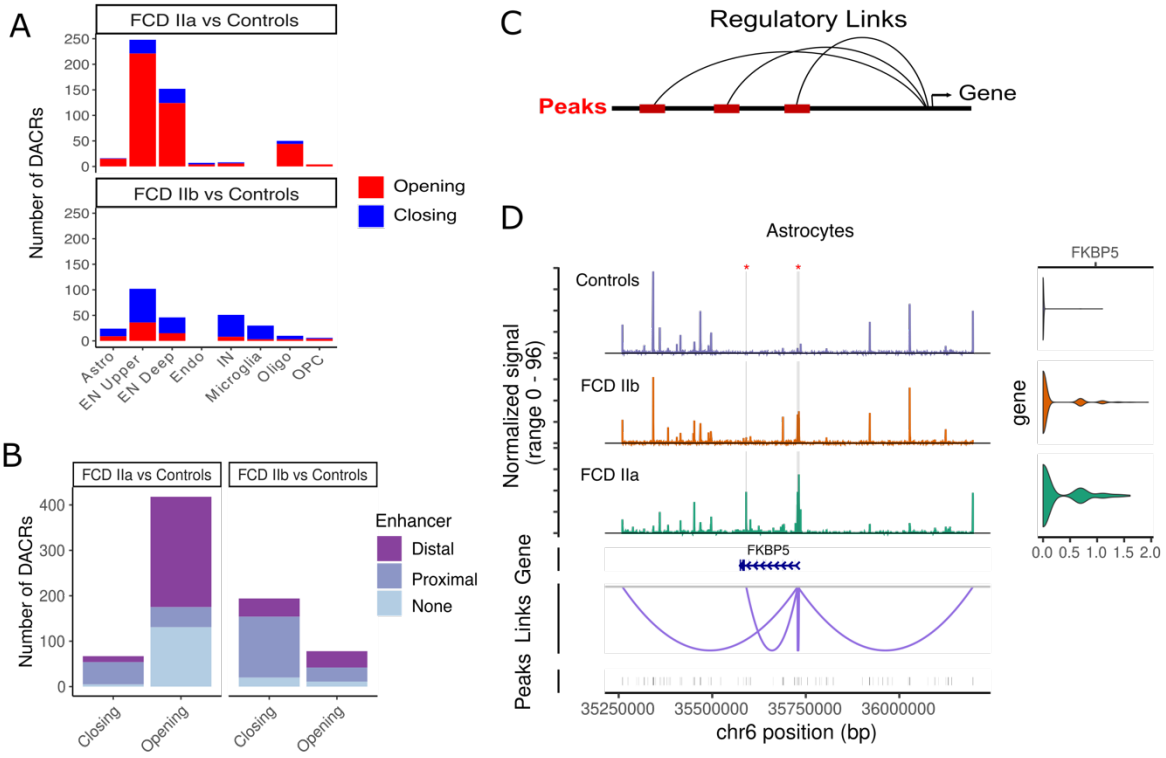


Figure 5. Open chromatin changes and cell-type specific regulatory links in FCD.

(A) Barplots depicting differentially accessible chromatin regions (DACRs) between FCD tissues and controls. The number of chromatin regions with opening or closing status (gain or loss of chromatin accessibility, respectively) is indicated for each cell type.

(B) Enhancer types overlapping DACRs. The number of opening and closing peaks associated with distal or proximal enhancers from Encode are indicated.

(C) Diagram depicting the inference of regulatory links identifying neighboring open chromatin regions (ATAC peaks) associated with gene expression using linear regression (see Methods).

(D) Regulatory links predicted for *FKBP5* in astrocytes. ATAC coverage tracks and corresponding expression levels in controls, FCD IIb, and IIa tissues are depicted at the top. Regions highlighted in gray indicate opening peaks in lesions and red asterisks indicate Encode-annotated enhancers. The gene track shows the location of the *FKBP5* locus in chromosome 6, and the bottom track shows all peaks contained in the chromosomal region.

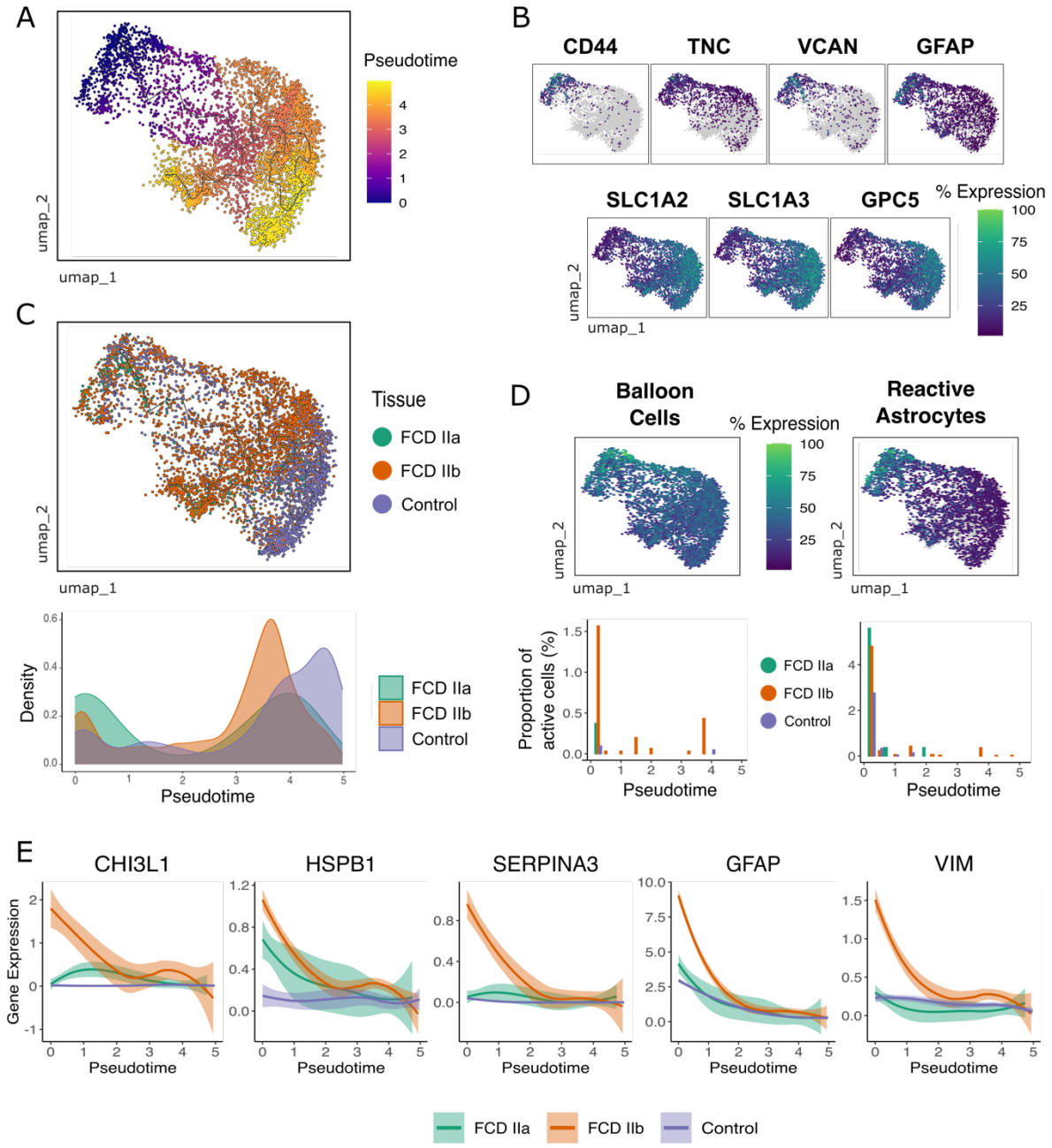


Figure 6. Impaired Astrocyte Differentiation in FCD.

(A) UMAP visualization of astrocytes temporally ordered on a trajectory as a function of pseudotime built by Monocle3.

(B) Feature plots depicting gene expression of immature astrocyte markers *CD44*, *TNC*, *VCAN*, and *GFAP* (top), and differentiated astrocyte markers *SLC1A2*, *SLC1A3*, and *GPC5* (bottom).

(C) UMAP visualization and density quantification along the trajectory. Density plots indicate the normalized density of cells along the pseudotime according to tissue condition.

(D) Feature plots depicting expression of balloon cells and reactive astrocytes gene signatures computed by AddModuleScore from Seurat (top). Bar graphs (bottom) quantify the proportion of active cells for these signatures along the pseudotime trajectory. Active cells were defined using the AUCell method and are colored by tissue condition.

(E) Expression levels of genes associated with balloon cells and reactive astrocytes *CHI3L1*, *HSPB1*, *SERPINA3*, *GFAP*, and *VIM* in the trajectory and by tissue condition.